Extraction of SIFT Features for Identifying Disaster Hit areas using Machine Learning Techniques

1st V. Suresh Kumar
Department of Information Technology
Vel Tech Multi Tech Dr. Rangarajan Dr. Sakunthala Engineering College
Chennai, India
sureshkumar@veltechmultitech.org

2nd Ahmed Alemran
Physics department
Medicine college-Misan university
Al Amarah, Iraq
ahmed.mcm@uomisan.edu.iq

3rd Shashi Kant Gupta
Department of CSE
Integral University
Lucknow, India
raj2008enator@gmail.com

4th Bramah Hazela
Computer Science & Engineering, Amity School of Engineering & Technology, Lucknow Amity University,
Uttar Pradesh, India
bramahhazela77@gmail.com

5th Chandra Kumar Dixit
Department of Physics
Dr. Shakuntala Misra National Rehabilitation University
Lucknow, India
ckparadise@gmail.com

6th Bhadrappa Haralayya
Department of MBA
Lingaraj Appa Engineering College Bidar
Bidar, India
bhadrappabhavimani@gmail.com

2022 International Conference on Knowledge Engineering and Communication Systems (ICKECS) | 978-1-6654-5637-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICKECS56523.2022.10060037

***Abstract*—Beyond the death toll, disasters have major and often devastating social and economic ramifications, including financial damage and loss. As a result of recent advances in artificial intelligence (AI), particularly in machine learning (ML) and deep learning (DL), disasters can be better managed. An emphasis has been placed on disaster and hazard prediction studies as well as assessments of risk and susceptibility and the detection of disasters as well as the implementation of early warning systems and disaster monitoring. A number of current disaster management applications have also been examined. The findings are discussed, and suggestions for further research are offered. As computing power rises and data availability expands, new flood detection methods are being developed using machine learning. To better understand nonlinear systems and make flood forecasts, machine learning has become a go-to tool. Hydrologic and hydraulic models that explain physical processes can be used in flood forecasting in conventional ways to predict hazard factors. Although these models may be used to gain a better understanding of a system, they frequently need a large amount of computing and data resources. Methods based on machine learning have the potential to enhance accuracy while also decreasing calculation time and model creation costs. As a result of this study, we now have a better understanding of the current status of flood forecasting using machine learning algorithms. This paper proposes a machine learning model for SIFT extraction. Once the extraction is done comparative analysis is done with the SVM and Fine Tree classifiers to prove the efficiency of the proposed model.**

Keywords— Flood, Disaster, SIFT features, Fine Tree, SVM, Machine Learning.

I. INTRODUCTION

Disaster management is a long-term approach to dealing with calamities. Mitigation, readiness, reaction, and recovery are now commonly accepted as four different stages. The term "mitigation" refers to efforts that may either prevent a catastrophe from occurring or lessen its effects. Emergency planning, stockpiling of supplies, and community education and training are all examples of preparedness efforts that help communities be better equipped to deal with disasters and their aftereffects. Protection of human life and property, natural resources, and the community's socioeconomic structure are all part of a community's response to an emergency [1]. A lot of attention is paid to this phase since it is the most critical for both people and infrastructure. The long-term efforts to restore the community to a condition of normality are part of recovery (reconstruction). Financial help, as well as rebuilding and reconstruction, all fall within this phase's purview (e.g., of buildings and key infrastructure). Community resilience may also be increased by including residents in catastrophe planning and response. Many ML and DL methodologies are utilized to aid disaster management in all stages. In order to design systems that can forecast catastrophes and aid with the reaction and recovery after such occurrences.

ML allows the use of large and complicated datasets and realistic decision-support tools to be built and used. Intelligence gained through manipulating disparate data sets and discovering patterns that would otherwise be inaccessible is the goal of these strategies. Large amounts of data may be gleaned from a variety of sources, including satellite photography. There have been recent disaster management evaluations focusing on AI disaster management applications, as well as the usage and promise of big data in natural disasters. The use of artificial intelligence (AI) to evaluate and interpret large amounts of data from multiple sources is on the rise in the disaster management industry. In all, 26 artificial intelligence (AI) approaches were employed in [2] for 17 disaster management application domains.

Discriminant machine learning aims to identify a discriminant function that can correctly predict labels for newly acquired instances, based on an independent and identically distributed training dataset [3]. Data points are assigned to one of the many classes in the classification task using discriminant classification functions, rather than generative machine learning approaches that use conditional probability distributions. When dealing with a multidimensional feature space and just needing posterior probabilities, discriminant approaches are a better choice for predicting outliers since they use fewer computer resources and training data. While learning a classifier, it's similar to figuring out the ideal equation for a multi-dimensional surface that best separates the various classes inside the feature space. Unlike genetic algorithms (GAs) and perceptron's, which are often used for classification in machine learning, SVM delivers the same ideal hyperplane parameter since it solves the convex optimization issue analytically. Perception problems may be solved in large part by considering the commencement and termination criteria.

978-1-6654-5637-1/22/$31.00 ©2022 IEEE

Authorized licensed use limited to: Indian Institute of Technology Patna. Downloaded on March 19,2023 at 14:35:40 UTC from IEEE Xplore. Restrictions apply.

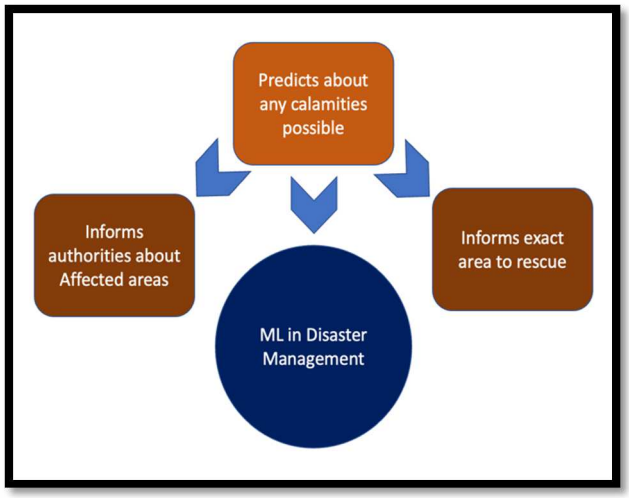

Predicts about any calamities possible
Informs authorities about Affected areas
Informs exact area to rescue
ML in Disaster Management

Fig. 1. Machine Learning Applications

Images are transformed into scale-independent coordinates based on local characteristics using the Scale Invariant Feature Transform (SIFT) [5]. Images are classified by Machine Learning (ML) to see the exact places in the dataset as shown in figure 1. Because of the way it creates so many details, the picture is densely covered across all sizes and locations. About 2000 stable characteristics may be found in a typical 500x500 pixel picture (although this number depends on both image content and choices for various parameters). Small items in a crowded backdrop need at least three elements to be accurately matched from each object for successful identification, hence the number of features is critical for object recognition. In order to match and recognize photos, a database of SIFT characteristics is first compiled from a collection of reference images [6]. Related works will give the existing work's work which is related to the ML Concepts. This proposed work has SVM and Fine Tree concept for training the data and test accordingly, also categorized the exact flood area with good accuracy value.

II. Related Works

Artificial intelligence (AI) and machine learning (ML) [7] have made it possible for buildings to learn and adapt to their environments. There has been a great deal of work done on machine learning applications for certain stages in the life cycle of a building. Real-world data clustering analysis issues cannot be handled by these techniques. Clustering methods need to be more versatile, efficient, and adaptable. Research and practice alike will benefit from this survey report. Transmissions over shorter lengths under 100 kilometers have adapted long-haul transmission methods for use in their shorter-distance counterparts. There are numerous problematic elements of optical communications that can be solved using ML approaches [9].

Unfortunately, surgical data scientists do not have access to video data sets that show adverse outcomes, which might bias models and make them less generalizable. Because blood obscures the picture, hemorrhage poses a special problem for computer vision–based models [10]. An annotated surgical video dataset of bleeding complication care, is accessible to the general public. Adaptation to catastrophes has been aided by recent advances in AI, particularly in machine learning and deep learning [11]. ML and DL created disaster management strategies are the topic of this study, which seeks to offer an overview of research findings released since 2017. Social media networks [12] like Twitter create enormous amounts of unstructured data that may be analyzed using text mining and machine learning algorithms. NLP and classification methods will be used in this study to identify actual and false catastrophe tweets. Tweets about genuine catastrophes and tweets about made-up disasters were included in the dataset, which was downloaded from the Kaggle website.

Policy and administrative choices [13] involving functional operations and emergency resources may be generated via disaster management. The disaster management cycle consists of four stages: prevention, planning, reaction, and recovery. Technology advancements have increased the pace at which IoTs and sensor devices generate disaster-related data. Disaster mitigation, preparation, response, and recovery may be aided by the availability of such data. Drought is a difficult natural catastrophe to anticipate because of its complexity and severe effects. Machine learning techniques (MLMs) [14] and their potential for use in the development of effective drought forecasting models are the subject of our study. Since then, MLMs' algorithms for drought modelling have seen substantial improvements in terms of resilience, efficacy, and accuracy. The Pima Indians Diabetes Dataset is used to develop an algorithm for diabetes categorization in pregnant women (PIDD) [15]. TFKNN, a fuzzy KNN built on uncertainty classifiers and adjusted to accommodate belief degrees, manages membership functions and operation law, all while avoiding incorrect categorization and misclassifications. For accuracy, specificity, precision, and AUC, TFKNN beat the others in every test.

The purpose of this study is to aid agriculture in maintaining a steady supply of food. Automatic classification of seeds allows researchers to keep tabs on the health of rice plants' seed crops, resulting in high-quality postproduction. Images of healthy and ill rice plants were processed to extract the red, green, and blue (RGB) [16] color characteristic in the seed's pixel point. Thus, in this experiment, the machine-learning method was used, since the classifier had attained a 100% accuracy and training time of 0.32189 seconds. This year, Machine Learning's [17] influence and utilization have grown significantly. A significant part of the Machine Learning Lifecycle is Algorithm Selection. Using a pre-processed dataset, this study compares a variety of machine learning techniques. With the help of an innovative method, it also aspires to supply the optimal hyper-parameters for the chosen algorithm itself. Using the Bayesian Neural Network classifier [18] with backtracking, a hierarchical learning approach is suggested to facilitate large-scale picture categorization. With a Visual Confusion Label Tree, it is possible to organize enormous picture collections into hierarchical structures. The tree classifier's classification skill improves because it focuses less on the complete set of categories and more on the subsets that cause confusion.

The procedures of recognizing a face and determining a person's name and gender are both difficult. A number of machine learning approaches were utilized. LBPH's face recognition accuracy was greater than those of other studies, and its gender identification accuracy was extremely near to that of other relevant studies, according to the present and previous research's results. Using methods like as SIFT, SURF, and LBP, this technology, which is in high demand right now, extracts characteristics from an input picture and compares them to those in another image to validate or assert an individual's identification. SVM classifier [20] – [22]

Authorized licensed use limited to: Indian Institute of Technology Patna. Downloaded on March 19,2023 at 14:35:40 UTC from IEEE Xplore. Restrictions apply.

coupled with SIFT approach surpasses other classifiers when compared. Processes based on data must make room for AI-driven frameworks [23]. The development of technology is focused on reducing the amount of data that must be transmitted and the amount of network traffic that must be handled [24].

III. METHODOLOGY

We have built a machine learning model for disaster hot areas. In this study we have taken into account the flood-hit areas. The machine needs to classify the whether the given area is flood or not. The overall workflow of the proposed model is shown in Fig. 2.

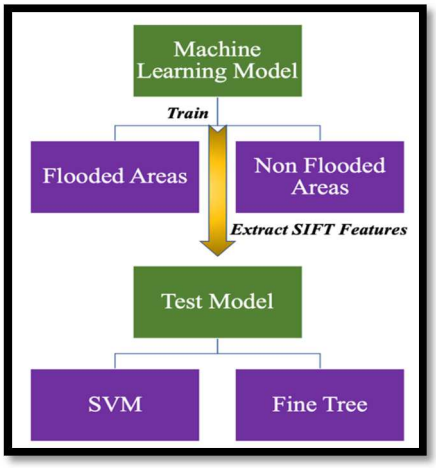


Fig. 2. Overall Workflow

The model needs to be trained with plenty of flooded and non-flooded images. The SIFT features are extracted from the images while getting trained. Once the features are extracted, the model is known with how the features of flood-affected area will be and how it would be for non-flooded areas. Now the model needs to be tested by giving the input as the test images. Once the images are fed into the classifier, the model is now capable to distinguishing the flooded areas with non-flooded areas as shown in Fig. 3 & 4.

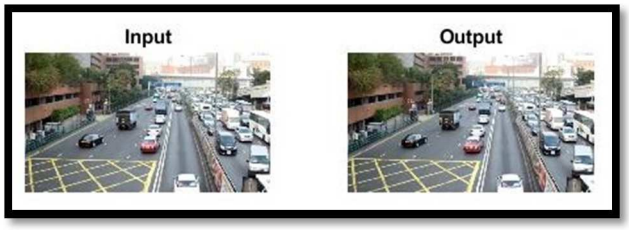
(a)

(b)

Fig. 3. Classification of flooded areas

Once the classifier is able to identify the flooded areas, we can also test the model by giving images that are not flood affected and observe the results. With these images also the classifier is able to well detect that there is no occurrence of flood for the given image as shown in Fig. 5. Specific interest spots in a picture can be identified using SIFT characteristics that have no effect on the image's size or rotation. Despite changes in light, noise, and even tiny shifts in perspective, they remain stable. Thus by extraction of these SIFT feature the model is able to identify the exact flooded areas easily in a very short time. Use of Fine Tree and SVM is done to obtain the various experimental result while working with the proposed SIFT feature extraction method.

(a)



(b)

Fig. 4. Non-flooded areas

IV. RESULTS AND DISCUSSION

The classifiers were tested using MatLab R2017B version. The input data was given both to Fine Tree and SVM classifier. This was done to examine the results obtained while using two classifiers for the same proposed extraction method. The confusion matrix for Model 1 that is the Fine Tree is shown in Fig. 5. From the matrix it can be observed that the true positive rate while using the FIne Tree classifier was about 97% and the false negative rates was about 3%. The ROC curve for the same was also computed and is shown in the Fig. 6. The same was done with the SVM classifier. The confusion matrix for the SVM classifier is shown in Fig. 7. From the matrix is could be observed that the True positive rate of the model was about 99% with a false negative rate of only 1%. The ROC curve of the SVM model is depicted in Fig. 8. The results obtained for both the classifiers are charted in Table 1. From the table it is evident that SVM performs better while comparing with the Fine Tree classifier. This is because of the linearity of the classifier. The execution time while using the SVM classifier was also less while applying it to the proposed SIFT based feature extraction method for classifying flood hit areas.

Authorized licensed use limited to: Indian Institute of Technology Patna. Downloaded on March 19,2023 at 14:35:40 UTC from IEEE Xplore. Restrictions apply.

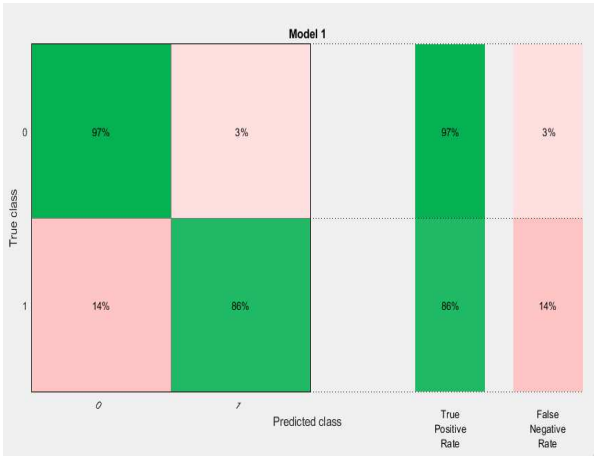


Fig. 5. Confusion Matrix - Fine Tree

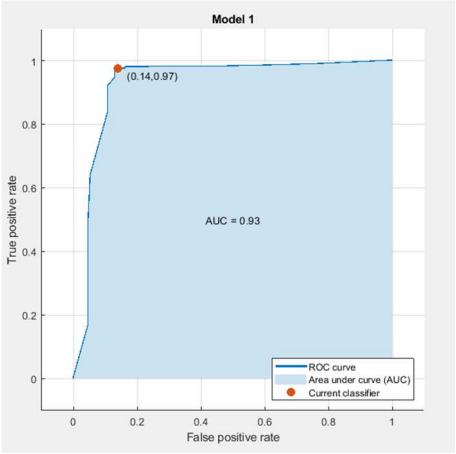


Fig. 6. RoC Curve - Fine Tree

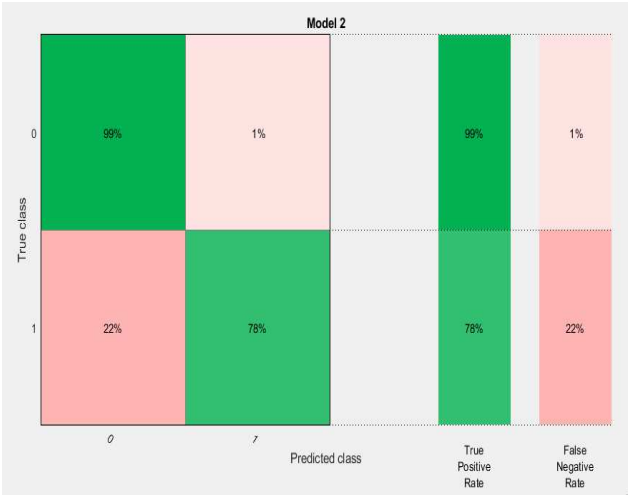


Fig. 7. Confusion Matrix - SVM

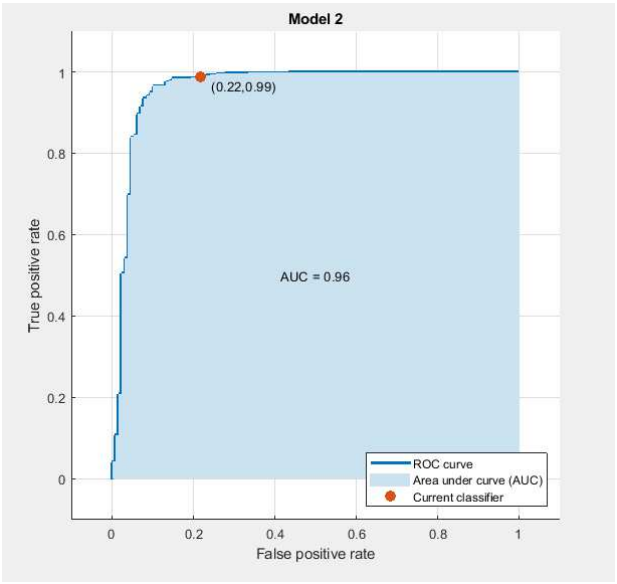


Fig. 8. RoC Curve - SVM

TABLE I. COMPARISON RESULTS - FINE TREE VS SVM

Model	Fine Tree	SVM
Accuracy	95.6%	95.3%
Prediction Speed	410obs/sec	350obs/sec
Training Time	44.276sec	31.946sec

V. CONCLUSIONS

Floods are some of the most destructive and hard to predict natural disasters. The research on improving flood prediction models helped reduce risks, make policy suggestions, reduce the number of people who died in floods, and lessen the damage to property caused by floods. In the past 20 years, machine learning (ML) methods have helped a lot to improve prediction systems by making them more like the complex mathematical expressions of the physical processes of floods. This paper suggests a model for SIFT extraction that is based on machine learning. Once the extraction is done, the SVM and Fine Tree classifiers are compared to the proposed model to show how well it works. With the experimental results obtained it could be claimed that the SVM classifier works better than the Fine Tree classifier for the proposed model of SIFT based feature extraction machine learning model.

REFERENCES

[1] Mosavi, A., Ozturk, P., & Chau, K. W. (2018). Flood prediction using machine learning models: Literature review. Water, 10(11), 1536.

[2] Akshya, J., & Priyadarsini, P. L. K. (2020). A Comparison of Machine Learning Approaches for Classifying Flood-Hit Areas in Aerial Images. In International Conference on Innovative Computing and Communications (pp. 407-415). Springer, Singapore.

[3] de Bruijn, J. A., de Moel, H., Jongman, B., de Ruiter, M. C., Wagemaker, J., & Aerts, J. C. (2019). A global database of historic and real-time flood events based on social media. Scientific data, 6(1), 1-12.

[4] Akshya, J., & Priyadarsini, P. L. K. (2019, February). A hybrid machine learning approach for classifying aerial images of flood-hit areas. In 2019 International Conference on Computational Intelligence in Data Science (ICCIDS) (pp. 1-5). IEEE.

[5] Pereira, J., Monteiro, J., Estima, J., & Martins, B. (2019, November). Assessing flood severity from georeferenced photos. In Proceedings of the 13th workshop on geographic information retrieval (pp. 1-10).

[6] Pereira, J., Monteiro, J., Silva, J., Estima, J., & Martins, B. (2020). Assessing flood severity from crowdsourced social media photos with deep neural networks. Multimedia Tools and Applications, 79(35), 26197-26223.

Authorized licensed use limited to: Indian Institute of Technology Patna. Downloaded on March 19,2023 at 14:35:40 UTC from IEEE Xplore. Restrictions apply.

[7] Alanne, K., & Sierla, S. (2022). An overview of machine learning applications for smart buildings. *Sustainable Cities and Society*, *76*, 103445.

[8] Ezugwu, A. E., Ikotun, A. M., Oyelade, O. O., Abualigah, L., Agushaka, J. O., Eke, C. I., & Akinyelu, A. A. (2022). A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects. *Engineering Applications of Artificial Intelligence*, *110*, 104743.

[9] Xie, Y., Wang, Y., Kandeepan, S., & Wang, K. (2022, January). Machine learning applications for short reach optical communication. In *Photonics* (Vol. 9, No. 1, p. 30). MDPI.

[10] Kugener, Guillaume, Dhiraj J. Pangal, Tyler Cardinal, Casey Collet, Elizabeth Lechtholz-Zey, Sasha Lasky, Shivani Sundaram et al. "Utility of the simulated outcomes following carotid artery laceration video data set for machine learning applications." *JAMA network open* 5, no. 3 (2022): e223177-e223177.

[11] Linardos, V., Drakaki, M., Tzionas, P., & Karnavas, Y. L. (2022). Machine Learning in Disaster Management: Recent Developments in Methods and Applications. *Machine Learning and Knowledge Extraction*, *4*(2), 446-473.

[12] Alhammadi, H. (2022). Using Machine Learning in Disaster Tweets Classification.

[13] Shukla, S., & Kumar, T. V. (2022). Machine Learning for Disaster Management. *AIJR Abstracts*, 55.

[14] Prodhan, F. A., Zhang, J., Hasan, S. S., Sharma, T. P. P., & Mohana, H. P. (2022). A review of machine learning methods for drought hazard monitoring and forecasting: Current research trends, challenges, and future research directions. *Environmental Modelling & Software*, 105327.

[15] Salem, H., Shams, M. Y., Elzeki, O. M., Abd Elfattah, M., F. Al-Amri, J., & Elnazer, S. (2022). Fine-tuning fuzzy KNN classifier based on uncertainty membership for the medical diagnosis of diabetes. *Applied Sciences*, *12*(3), 950.

[16] Geollegue, K. W. V., Arboleda, E. R., & Dizon, A. A. (2022). Seed of rice plant classification using coarse tree classifier. *Int J Artif Intell*, *11*(2), 727-735.

[17] Avinash, M., Nithya, M., & Aravind, S. (2022, May). Automated Machine Learning-Algorithm Selection with Fine-Tuned Parameters. In *2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 1175-1180). IEEE.

[18] Liu, Y., Dou, Y., Jin, R., Li, R., & Qiao, P. (2022). Hierarchical learning with backtracking algorithm based on the visual confusion label tree for large-scale image classification. *The Visual Computer*, *38*(3), 897-917.

[19] Alamri, H., Alshanbari, E., Alotaibi, S., & AlGhamdi, M. (2022). Face Recognition and Gender Detection Using SIFT Feature Extraction, LBPH, and SVM. *Engineering, Technology & Applied Science Research*, *12*(2), 8296-8299.

[20] Dorňák, V., Čermák, M., Pecha, M., & Pospíšil, L. (2022, April). SIFT feature extraction applied in SVM classification. In *AIP Conference Proceedings* (Vol. 2425, No. 1, p. 170002). AIP Publishing LLC.

[21] Hicham, B., Ahmed, C., & Abdelali, L. (2022). Face recognition method combining SVM machine learning and scale invariant feature transform. In *E3S Web of Conferences* (Vol. 351, p. 01033). EDP Sciences.

[22] Bilik, S., & Horak, K. (2022). SIFT and SURF based feature extraction for the anomaly detection. *arXiv preprint arXiv:2203.13068*.

[23] Navaneetha Krishnan Rajagopal, Naila Iqbal Qureshi, S. Durga, Edwin Hernan Ramirez Asis, Rosario Mercedes Huerta Soto, Shashi Kant Gupta, S. Deepak, "Future of Business Culture: An Artificial Intelligence-Driven Digital Framework for Organization Decision-Making Process", Complexity, vol. 2022, Article ID 7796507, 14 pages, 2022. https://doi.org/10.1155/2022/7796507

[24] Eshrag Refaee, Shabana Parveen, Khan Mohamed Jarina Begum, Fatima Parveen, M. Chithik Raja, Shashi Kant Gupta, Santhosh Krishnan, "Secure and Scalable Healthcare Data Transmission in IoT Based on Optimized Routing Protocols for Mobile Computing Applications", Wireless Communications and Mobile Computing, vol. 2022, Article ID 5665408, 12 pages, 2022. https://doi.org/10.1155/2022/5665408

Authorized licensed use limited to: Indian Institute of Technology Patna. Downloaded on March 19,2023 at 14:35:40 UTC from IEEE Xplore. Restrictions apply.